

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 11:50:52 ; Search time 2342.25 Seconds
(without alignments)
5151.576 Million cell updates/sec

Title: US-09-788-476a-1
Perfect score: 894
Sequence: 1 gggagtgagtgagtgagtgtaa.....taataaaaaaatagaana 894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_iny:*
15: em_gss_pln:*
16: em_gss_vtr:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	839	93.8	910	11	AF161434	AF161434 Homo sapi
2	821	1070	910	11	BM460786	BM460786 AECNCOURT
3	780.6	87.3	806	10	BG574651	BG574651 602596653
4	765.2	85.6	788	10	BM462456	BM462456 AECNCOURT
5	759	84.9	783	10	BG910161	BG910161 602805548
6	749.4	83.8	783	10	BG533012	BG533012 602580893
7	741.4	82.9	887	10	BI255433	BI255433 602977928
8	740.2	82.8	776	10	BI460375	BI460375 603202043
9	733	82.0	889	10	BI552670	BI552670 603193915
10	725.4	81.1	774	10	BI870282	BI870282 603393788
11	723	80.9	717	10	BI520421	BI520421 603071604
12	716	80.1	744	10	BG493148	BG493148 602541829
13	714.6	79.9	781	10	BG720065	BG720065 602691562
14	711.8	79.6	974	10	BG387900	BG387900 602412920
15	705.8	78.9	1139	10	BG289677	BG289677 602384642
16	697.2	78.0	802	10	BG542056	BG542056 602572169
17	663.8	74.3	1107	10	BG250284	BG250284 602362415

18	662.4	74.1	699	10	BG261221	BG261221 602372957
19	652.6	73.0	781	10	BG178461	BG178461 602330263
20	651	72.8	685	10	BG390201	BG390201 602415736
21	647	72.4	648	9	AV710050	AV710050 AV710050
22	640	71.6	799	10	BG500895	BG500895 602547362
23	639.6	71.5	747	10	BE391455	BE391455 601286447
24	635	71.0	718	10	BG502106	BG502106 602548449
25	629.4	70.4	798	10	BG323390	BG323390 602421786
26	622.8	69.7	897	10	BF576251	BF576251 602134108
27	613.8	68.7	666	10	BI460158	BI460158 603201744
28	612.4	68.5	683	10	BF092754	BF092754 MR4-TN011
29	606	67.8	610	10	BG614223	BG614223 602639906
30	602.4	67.4	933	10	BG288241	BG288241 602383764
31	598.6	67.0	1107	10	BF979954	BF979954 602288518
32	596.8	66.8	893	11	AK003453	AK003453 Mus muscu
33	595.2	66.6	754	10	BF683147	BF683147 602139209
34	594.2	66.5	910	10	BE378791	BE378791 601237288
35	579.2	64.8	699	10	BI961456	BI961456 MONOL 5.G
36	577	64.5	681	10	BE378430	BE378430 601236768
37	574.6	64.3	773	10	BE782311	BE782311 601470086
38	574.4	64.3	577	9	AM131598	AM131598 xE31h09.x
39	571	63.9	571	10	BE887511	BE887511 601508160
40	559.8	62.6	639	10	BE32774	BE32774 601307843
41	557	62.3	784	10	BG254640	BG254640 602368607
42	556	62.2	631	10	BE390709	BE390709 601286887
43	551.2	61.7	624	10	BF888562	BF888562 MR4-TN011
44	551	61.6	562	9	BE018801	BE018801 bbs5b11.y
45	547.8	61.3	826	10	BG196492	BG196492 RST15716

ALIGNMENTS

RESULT	1	AF161434	910 bp	MRNA	linear	HTC 22-MAY-2001
LOCUS	AF161434	Homo sapiens	HSPC316 mRNA, partial cds.			
DEFINITION	AF161434	Homo sapiens	HSPC316 mRNA, partial cds.			
ACCESSION	AF161434	AF161434.1	GI:6841281			
VERSION	AF161434.1	GI:6841281				
KEYWORDS	HTC.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.					
AUTHORS	Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L., Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J. and Chen,Z.					
TITLE	Human partial CDS from cd34+ stem cells					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 910)					
AUTHORS	Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L., Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J. and Chen,Z.					
TITLE	Direct Submission					
JOURNAL	Submitted (14-MAY-1999) Shanghai Institute of Hematology, Shanghai Second Medical University, Rui-Jin Hospital, 197 Rui-Jin Road II, Shanghai 200025, P. R. China					
FEATURES	source					
source	Location/Qualifiers					
	1..910					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="CBIAIH07"					
	/cell_type="cd34+ stem cell"					
	/tissue_type="blood"					
	<1..769					
	/codon_start=-2					
	/product="HSPC316"					
	/protein_id="AAF28994.1"					
	/db_xref="GI:6841282"					
	/translation="SGYRKKMATEYELKRLAEKQCECLARGLETGKIDDLIR					
	LOAVLEHAEAEANEDVLDDETEETKRIEIVPEKEEPEPTVDAARKKVITIS					
	ELIOTERKRAERFNFVSLIESKARARAVGISVFNKRSVLLINMLTWISLKE					
	ELKDLGLNVSISIRKSEDDKLRKRERFCIVTSAGTGTTEARGRKEGSAIGLP					
	DEKFLISVLCFPFLSFVLTVPKCTVCLRPASQ"					

BASE COUNT 299 a 158 c 235 g 218 t
ORIGIN

Query Match 93.8%; Score 839; DB 11; Length 910;
Best Local Similarity 99.0%; Pred. No. 5e-153;
Matches 886; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```
Oy 3 gagtgaagtgagggtaaacagaatgagcagcagagcgtgtgagctccataaagct 62
    |||||||
Db 1 GAGTGGAGTGAAGGGGTAACAGATGGCGACGACAGCGTGGAGCTCCATTAAGCT 60
Oy 63 tgcgaactaaagcaagaatgtctgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 122
    |||||||
Db 61 TCCCGAACAAGCAAGAAATGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Oy 123 tcttatccagacacccagcagcagcagcagcagcagcagcagcagcagcagcagc 182
    |||||||
Db 121 TCTTATCCACAGACTCCAGGCAATCTTGAAGAACTGCTGAAGAGGCAATGAAGA 180
Oy 183 agatgtactgtgagatgaacacagagagagagagagagagagagagagagagag 242
    |||||||
Db 181 AGATGACTGGAGATGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Oy 243 agaggaagaagacccctgaaaaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgt 302
    |||||||
Db 241 AGAGGAAGAAACCCCTGAAAAAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Oy 303 taccctctgaataacacacagacagcagagagagagagagagagagagagagag 362
    |||||||
Db 301 TACATCTGAATTAACACAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Oy 363 tgtgtgagctgtgagatgaagaagctgtcgtgagcagcagcagcagcagcagcag 421
    |||||||
Db 361 TGTGAGCTTGGAGAGTAAGAAAGCTGCTGGGAGCTAGGAGTGGAGTTCCTTCA 420
Oy 422 caacaaaaggtctgtcat-c-tgtatacaaaacctatgtgttaactgtgataagc 480
    |||||||
Db 421 CAACAAAAGGTCTGTCAATCTGTCAATCTGTCAATCTGTCAATCTGTCAATCTGT 480
Oy 481 agagctcaagaagattc-ggttgaatgtctctcaatctcagaagagctcagaatga 539
    |||||||
Db 481 AGAGCTCAAGATTTGGGTTGTAATGTCTTCAATCTGTCAATCTGTCAATCTGT 540
Oy 540 gaacatgaaaaaagagagagagagagagagagagagagagagagagagagagag 599
    |||||||
Db 541 GAAACTGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Oy 600 cagcagaagatcacagaagcagaagaagaagaagaagaagaagaagaagaagaaga 659
    |||||||
Db 601 CACAGAGGATACAGAGGC-AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
Oy 660 aaaagctccagtaactctgtctccagtgcttccatcttccctcctcctgtgta 719
    |||||||
Db 660 AAAAGTCTGTACTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTGT 719
Oy 720 catatagcctaataatgacagatcagtgagcagcagcagcagcagcagcagcagc 779
    |||||||
Db 720 CATATATGCTTAATATGACAGATGATGCTACGTCTGTCTGTCTGTCTGTCTGT 779
Oy 780 accccaagtgatacgaatgacagcagcagcagcagcagcagcagcagcagcagc 839
    |||||||
Db 780 ACCCCAGGATACATCATGAACTGCGGACACAGCTTGAATTTAGCTTGAAG 839
Oy 840 gttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 894
    |||||||
Db 840 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 894
```

RESULT 2
LOCUS BM460786 1070 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6421520 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532141

5', mRNA sequence.
BM460786
EST.
GI:18509826

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1070)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLML12214 row: 0 column: 22
High quality sequence stop: 593.

FEATURES
source
location/Qualifiers
1..1070
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5532141"
/clone_image="NIH_MGC_71"
/issue_type="telomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; site_1: NotI;
site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

BASE COUNT 317 a 231 c 278 g 237 t 7 others

ORIGIN

Query Match 91.8%; Score 821; DB 10; Length 1070;
Best Local Similarity 98.2%; Pred. No. 1.5e-149;
Matches 861; Conservative 0; Mismatches 12; Indels 4; Gaps 3;

```
Oy 15 gggttaacaagatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 74
    |||||||
Db 4 GGGGAACAAGATGGCCAGCAGAGCTGAGCTCCATTAAGCTTAAGCTTCCGAACTAA 63
Oy 75 ggaagaatgtctgtcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 134
    |||||||
Db 64 GCAGAGATGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 123
Oy 135 actccagcatalcttgaagaacatgtctgaagaagagagagagagagagagagagag 194
    |||||||
Db 124 ACTCCAGCATATCTTGAAGAACATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 183
Oy 195 agatgaagaagagagagagagagagagagagagagagagagagagagagagagag 254
    |||||||
Db 184 AGATGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
Oy 255 cccctgaanaaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 314
    |||||||
Db 244 CCTGTAANAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 303
Oy 315 acccagacatgagagagatcagaagaagagcagcagcagcagcagcagcagcagcag 374
    |||||||
Db 304 ACCACAGATGAGAGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 363
Oy 375 gggtlaagaagctgtcgtgagcagcagcagcagcagcagcagcagcagcagcagcag 434
    |||||||
Db 364 GAGTAAGAAAGCTGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
Oy 435 gtcactctgatacaaacctatgtgttaactgtgataagcagagagagagagagagag 494
    |||||||
Db 424 GTCACTGTGATTAACAACCTATGATTAAGCTGAAGAGAGAGAGAGAGAGAGAGAT 483
```

BASE COUNT	ORIGIN
263 a	155 c 211 g 177 t

Query Match	87.3%	Score 780.6	DB 10	Length 806
Best Local Similarity	99.4%	Pred. No. 1.1e-141		
Matches 794	Conservative 0	Mismatches 4	Indels 1	Gaps 1
QY 3	gagtgagtgaggggtaacaaagatgagccagagccggttgagagctccctaagctaaact	62		
Db 8	GGGTCCGGTGAAGGGGTTAACAGATGGCCGACGAGCGGGTGGAGCTCCATTAAAGCTAAAGCT	67		
QY 63	tgccgaactaaagcaagaatgtctgtctcgttggtttgagaccagagggaataaagcaaga	122		
Db 68	TCCCGAAGCTAAAGCAAGATGTCGTCTGCTGTGTTGGAGACCAAGGGGATTAACACAGA	127		
QY 123	tcttatccacagactccagcatctcttgaagaacatctgaaggagcgcaatlgaaaga	182		
Db 128	TCTTATCCACACATCCACAGGCGATATCTTGAGAACATCTGTAAAGAGGAGCCAAATGAAGA	187		
QY 183	agatttactgggagatlgtaaacagaggaaagaaagaaacaaagcccatggagctccctgttaa	242		
Db 188	ACATGTACTGGGAGATGAACACGAGGAAGAAACAAACAAAGCCCATTTAGCTCCGTGTA	247		
QY 243	agaaggaagaacccccctgtaaaaaaacctgttgaatgtggcagcagagaaagaaagtggtgaaat	302		
Db 248	ACAGGAGAACCCCTCGTAAAAAATCGTTGATGTGGCACACAGAGAAAGAAAGTGTAAT	307		
QY 303	tacatctgaataaccacagacttgagagaatgcagaagaagggtctgaacatcaatgtacc	362		
Db 308	TTCATCTGTAATATCACACACAGACTGAGAGATGACAGAAAGGGCTGTAACCATTCATCAATGTACC	367		
QY 363	tgtgagcttggagagtaagaaagctgtgtccggcagctagagtttggagatttcttaagttcc	422		
Db 368	TGTGAGCTTTGGAGAGTAAAGAAAGCTGTCTGGGCAAGCTTAGGTTTGGGATTTTCTTCAGTTCC	427		
QY 423	aaacaaaggtctgtcatctgataacaaacctaagtttaacttgaataagctgtgaaggaag	482		
Db 428	AACAAAGGCTGTGTCATCTGATTAACAAACCTATGTATGTTAACTTGATGATGTAAGCAAGAAAG	487		
QY 483	agcttaagaatttggtttggaatgtctcttcaatctccagaaagttcgaatgatgatgaa	542		
Db 488	ACCTTAAGATTTGGTTTGAATGTCTTTCATCTCCAGAAAGTCTGTAAGATGATGAGAA	547		
QY 543	actgtaaaaagaggaagagcgatgttggagattgtacaagtctcaagcttggaacttggaaccac	602		
Db 548	ACTGAAAAAGAGAGAGAGACCGCATTTGGGATTTGTCAAAAGTTGACGTGGAACTGGAAACAC	607		
QY 603	agaagatacagaagcaaaagaaagaggaagaaagcagagcgtcttggagattgctgataa	662		
Db 608	AGAGATATACAGAGCAAAAGAAAGAGAAAAAGACAGAGCGCTTTGGGATTTGCCTGATGAAA	667		
QY 663	agttcctgatacttctgtcttccagtgatttcaatctctctctctctcttctgtgtcaat	722		
Db 668	ACTTCTCTGATATTCTCTGTCTCCAGTGTCTTCCATTTCTCTCTCTCTCTCTTGGTGGACAT	727		
QY 723	ataatgcctaaatgacacagatcatgtgtccctacgttccgtccgc-aatgagggagcatgtac	781		
Db 728	ATATGCGCTAAATGACACAGATCATGTGCTGCTTAACGTCCTCGCAATATGAGGAGCATGTAC	787		
QY 782	cccaggtacatcatgaac 800			
Db 788	CCCAGGTACATCCATGAAC 806			
RESULT 4				
BM462456				
LOCUS		788 bp	mRNA	linear
DEFINITION	AGNCCOURT_6426392 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5518124			
ACCESSION	BM462456			
VERSION	BM462456.1 GI:18511496			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Euteleostomi; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			


```

Db 121 CCAGCATATCTTGAAGACATGCTGAAGAGAGCCAAATGAAAGATGACTGGAGA 180
Qy 198 tgaacagaggaagaagaacaagcccatgagctccctgtcaagaaggaaaccccc 257
Db 181 TGAACAGAGAGAACAGAAACAAAGCCCATTTAGAGTCCCTGTCAAGAGAGAAACCCC 240
Qy 258 tgaacaaactgtatgtgacagagagaagaagtgtgtgaataatcatctgaatacc 317
Db 241 TGAACAAACCTGTATGTGCGCAGAGAAAGAAAGTGTGAAATTCATCATCTGAATACC 300
Qy 318 acaacagagagaaatgacagaagaagggtcgaacgattcaatgtactgtgagcttggagag 377
Db 301 ACACAGCTGAGGAATGCGAAGAGAGGCTGAAGCATCAATGATCTGAGCTTGGAGAG 360
Qy 378 taagaagagctgtcgggacagctagtttggatttcttcagttccaacaagaagctgtc 437
Db 361 TAAAGAGCTCTCGGGAGAGTAGTGGATTTCCTTCACTCCAAACAAAGGCTGTCTC 420
Qy 438 atctgatacaaacctatgttgaacttggatlaagctggaagaaagctcaagaatttg 497
Db 421 ATGTGATTAACAAACCTATGTTAGTTAACTTGATGAGCTGAAGAAAGAGCTCAAGATTGG 480
Qy 498 ttgaagtgctcttcaatctccagaagctcgaagatgatagagaactggaagaagagaa 557
Db 481 TTTGAATGCTCTTCAATCTCCAGAAAGCTGGAAGATGATGAGAAACTGAAAGAGAGAA 540
Qy 558 ggaacgatttggagattgtcacaagctcagctggaactggaacccacagagagatacagagc 617
Db 541 GGAGCGATTGGGATTTGTCACAAAGTTCAGCTGGAAGCTGGAACCCAGAGAGATACAGAGC 600
Qy 618 aaagaagaagaaagagacagacgcttggagctgagctgagc-aaaaattcctgatactt 676
Db 601 AAAAGAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 677 tctgttccagtgctt-ccatttctccttcttcttctgtgtacataatgacctaaatg 735
Db 661 TCTGTTCCTCCAGTGTTCCTCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Qy 736 cagagctatgctgacgctcctgctgacatgagagagagagagagagagagagagagag 795
Db 721 CACAGTCAATGCTGCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 796 tga 798
Db 781 TGA 783

RESULT 6
Bg533012 783 bp mRNA linear EST 03-APR-2001
LOCUS 602580893F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4718711 5',
DEFINITION mRNA sequence.
ACCESSION Bg533012
VERSION Bg533012.1 GI:13524551
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 783)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/MLNL at:
http://image.llnl.gov
Plate: L1CM1570 row: j column: 24

```

```

FEATURES
    source
        High quality sequence stop: 763.
        Location/Qualifiers
            1..783
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4718711"
            /clone_id="NIH_MGC_61"
            /tissue_type="embryonal carcinoma"
            /lab_host="DH10B (TI1 phage-resistant)"
            /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
            SfiI (ggccgctggcc); Site_2: SfiI (ggccattggcc);
            Double-stranded cDNA was prepared from cell line RNA. 5'
            and 3' adaptors were used in cloning as follows: 5'
            adaptor sequence: 5'-CACGCCATATGAGCC-3' and 3' adaptor
            sequence: 5'-ATCTAGAGCGCGAGGCGCGGAGCAGTg-dh(30)BN-3'
            (where B = A, C, or G and N = A, C, G, or T). Average
            insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
            contained inserts by PCR. This library was enriched for
            full-length clones and was constructed by Clontech
            Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
            Library."
BASE COUNT      257 a      145 c      207 g      174 t
ORIGIN
Query Match      83.8%; Score 749.4; DB 10; Length 783;
Best Local Similarity 98.7%; Pred. No. 1.2e-135;
Matches 777; Conservative 0; Mismatches 6; Indels 4; Gaps 2;
Qy 12 gagggtaacaagaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 71
Db 1 GAGGGTAAACAAGATGGCGACCGAGAGCGTGGAGCTCATAGCTTAAGCTTGGCGAGCT 60
Qy 72 aaagaagaatgcttgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 131
Db 61 AAAGCAAGAAATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 132 cagactccagcatatcttgaagaacatgctggaagaaggcaatggaagaagtgtact 191
Db 121 CAGACTCCAGCATATCTTTGAAGAAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 192 gggagatgaacaagaagaagaagaacaagcccatgtagctccctgtcaagaagagaaga 251
Db 181 GGGAGATGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 252 accccctgaaaaaactgttgaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 311
Db 241 GCCCCTGAAAAAACTGTTGATGTGCGACAGAGAAAGAGTGTGAAATTTACATCTGA 300
Qy 312 aataccagagctgagagaatgacagaagaaggctgaacgattcaatgtactgtgagctt 371
Db 301 AATACCAAGACTGAGAAATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 372 ggaagatgaagaagctgtcggcagctaggttttggatttcttcagttccaacaagaag 431
Db 361 GGAGAGTGAAGAAAGCTGCTCGGGCAGCTAGGTTTGGAGTTCTTCAATTCACAAAGAG 420
Qy 432 tctgtcatctgatacaacaactatgttaacttggataagctggaagaagaagctcaag 491
Db 421 TCTGTCACTGTATTAACAAACCTATGTTAGCTTGAAGCTGAAGAGAAAGCTCAAG 480
Qy 492 atttggttgaatgtcttcaatctcagaagaagctgaagaatgataagctgaagaagct 551
Db 481 ATTGCTTGAATGCTCTTCAATCTCAGAAAGCTCAAGATGATGAGAAATTCATTA 540
Qy 552 gaggaagagcgaatttggatgtgtacaaagtcaagctggaactggaacacagagagatc 611
Db 541 GAGCAAGAGAGAGATTTGGATTGTCACAAGTTCACTGGAAGTCAAGAGAGAGATAC 600
Qy 612 agaggaagaagaagaagaagaagcagcagcttggatgtgtgtgtgtgtgtgtgtgtgtgt 671
Db 601 AGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659

```



```

|||||
Db 185 GGGAGATGAAACGAGAGAAAGAAACAAAGCCCATGTGCTCCCTGCTAAAGAGAA 244
Oy 252 accccctgaaaaaactgttgaatctgacagcaagaagaagtggtgaaataatcatctga 311
Db 245 ACCCCCTGAAAAAAGCTGTTGATGTGGCAGCAGAAAGAAAGTGCGTAAATATACATCTGA 304
Oy 312 aatcacacagactgagaagaatgcagaagaaggctgaacgatcaatgtacctgtgagctt 371
Db 305 AATACACAGACTGAGAGAAATGCAGAAAGGCGCTCAACATTCATATGTCCTGTGAGCTT 364
Oy 372 ggaagatgaagaagctgtctcggcgagctgaagtttgagatttcttcagttccacaagaag 431
Db 365 GGAGGATGAAGAAAGCTGCTGGGCGACTAGGTTGGGATTCTTCAGTTCCACAAAGAG 424
Oy 432 tctgcacactgaatacaaacctatgttaacttgataaagctgaaggaagaagctcaag 491
Db 425 TCTGTCACTGTAAACAAACCTATGTTTAACTTGATAGCTGAAGGAAGAGCTCAAG 484
Oy 492 atttggttgaatgtctctctcaatctcagaagaagctcgaagaatgatgaaactgaaaa 551
Db 485 AATTGGTTGAATGTCTCTTCATCTCCAGAAAGTCTGAAGATGATGAGAAACTGAAAA 544
Oy 552 gagaagaagagcatttggaattgtcaagaatgcagctgaagctgaagcaagaagaatgc 611
Db 545 GAGGAAGAGCATTGGGATTGTACACAGTTCACTGAGCTGGAACCTGACACAGAGATAC 604
Oy 612 agagcaaaagaaagaaagaaagcagcagcttgggaattgctgcatgaaaaagctctga 671
Db 605 AGAGGC -AAGAAGAGAAAGAAAGAGCAGAGCGCTTGGGATTCGCTGATGAAAAATCTCTGA 663
Oy 672 taacttctgtctcagtggttttccatctctctctctctctctctgttcacataatgctta 731
Db 664 TACTTTCTGTTCTCCAGTG -TTTCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 721
Oy 732 aatgacagctacatgtgcctgaagctcctgcctgcaatgaaggaagc -atgtaccacagga 789
Db 722 AATGCACACTGTGTGCTACGCTCTGCTCCCAATGAGGGAACCATGTTACCCCAAGTT 781
Oy 790 catcactgaactgcgcgcagctgtgaactatgtctgttccagctttaagttgtgtgt 849
Db 782 ACATCTTAAGTACGCGGAGCAG -TTGAATTAATTTGCTG -TTCAAGCTTAAGTGGAGAAAG 839
Oy 850 ttgtgtttgtatattgtgtgtgtgttaataaaaaaat 888
Db 840 GAGAGAAATGCTGCTGTATTAATAAATTTAAAAAAT 878

RESULT 10
LOCUS B1870282 774 bp mRNA linear EST 11-OCT-2001
DEFINITION 603393788F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5403909 5',
ACCESION B1870282
VERSION B1870282.1 GI:16043955
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 774)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMN at:
http://image.llnl.gov

```

```

Plate: LLM12029 row: p column: 22
High quality sequence stop: 774.
FEATURES
source
1..774
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5403909"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-Sport6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 254 a 145 c 201 g 173 t 1 others
ORIGIN
Query Match 81.1%; Score 725.4; DB 10; Length 774;
Best Local Similarity 98.6%; Pred. No. 5.6e-131;
Matches 763; Conservative 0; Mismatches 7; Indels 4; Gaps 3;
Oy 16 gataaagaatgagcagcagagcagcttgagctcctaagaactgaagcttgcgaactaaag 75
Db 1 GGTAAACAAAGTGGCGACGACGAGGAGGTGAGCTCCATTAAGCTTGCAGACTTAAAG 60
Oy 76 caagaatgctgtcgtgtgtgtgtgttgagacaaaggaagaataaagcaaatcttaaccacaga 135
Db 61 CAAGATGCTCTTGCTGCTGCTGCTTGGAGACCAAGGAATTAAGCATGCTTATCCACAGA 120
Oy 136 ctccagacatctctgaagaacatgctcgaagaagagcgaatgaagaagatgactgga 195
Db 121 CTCCAGGCTATCTTGAAACATGCTGAAGAGGAGGCGAAATGAAGATGTACTGGGA 180
Oy 196 gatgaacagagaagaagaagaagaagcccatctgagctccctgctcaagaagaagaagccc 255
Db 181 GATGAACAGAGAGAAAGAAACAAAGCCATTGAGCTCCCTGTCAAAGAGAAAGAACCC 240
Oy 256 cctgaaaaaacgttgaatgttgacagcagaagaagaagtgatgaataatcatctgaata 315
Db 241 CCTGAAAAAAGCTGTGATGTGCGACAGAGAAAGAAAGTGTGAATAATCATCTGAATTA 300
Oy 316 ccacagactgagaagaatgcagaagaagagcgtgaacgaattcaatgtacctgtgactgag 375
Db 301 CCACAGACTGAGAGAAATGCAGAAAGGCGCTGAACCATTCATGCTGAGCTTGGAG 360
Oy 376 agtaagaagcgtgtcgcgcgcagctaggtttggatttcttcagttccacaagaagctgt 435
Db 361 AGTAAGAAAGCTGCTCGGGCACTAGGTTTGGGATTTCTTCAGTTCCAAACAAAGGTTG 420
Oy 436 tcaatctgaatacaactatggtttaacttgaataagctgaagaagaagctcaaatatt 495
Db 421 TCATCTGATTAACAAACCTATGCTTAAGTAAAGCTGAAGCAAGAAAGAGCTCAAGATTT 480
Oy 496 ggtttgaatgtctcttcaatctccagaagaagcttgaagaatgaagaagaactgaagaagg 555
Db 481 GGTGTAATGTCTCTCAATCTCCAGAAAGTCTGAAGAGATGAAGAAATGTAAGAAAGAG 540
Oy 556 aaggaagcgaattgggattgtcaagaatgcagctggaactggaacacagagatacaag 615
Db 541 AAGGAGCGATTGGGATTTCACAAAGTTGAGTGAACCTGAAACCAAGAGATACAGAG 600
Oy 616 gcaagaagaagaaagaaagcagcagcgttgggaattgctgtcctgtga -aagctcctgatac 674
Db 601 GCAAGAAAGAGAAAGAGACAGAGCGCTTGGGATTCCTGTATGAACAAAGTCTCGATAC 660
Oy 675 ttctcg--ttctccagtggtttccatctctctctctc-ttgttcacataatgacta 731
Db 661 TTTCTGTTCTCTCAGTTGTTTCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
Oy 732 aatgacagctacatgtgcctgaagctcctgcctgcaatgaggaagcgtgaacca 785
|||||

```


DB	721	MATGCAACAGTATGTCCTACTGCTCCTGCGCATATGAGGAGCATGTACCCCA	774
RESULT	11		
LOCUS	B1520421		
DEFINITION	B1520421	817 bp	linear EST-29-AUG-2001
ACCESSION	60307160471 NIH_MGC_119	Human sapiens	cdna clone IMAGE:5163673 5',
VERSION	B1520421		
KEYWORDS	B1520421.1	GI:15345213	
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NIH-MGC http://mgs.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgsabrs@mail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMNL at:		
	http://image.llnl.gov		
	Plate: LMNL1406	row: 9	column: 02
	High quality sequence stop: 809.		
FEATURES			
source	Location/Qualifiers		
	1..817		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5163673"		
	/clone_lib="NIH_MGC_119"		
	/tissue_type="medulla"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: PCMV-SPORT6; Site_1: NotI;		
	Site_2: EcoRV (destroyed); RNA source normal medulla from		
	anonymous male age 27. Library is oligo-dT primed and		
	directionally cloned (EcoRV site is destroyed upon		
	cloning). Average insert size 1.3 kb, insert size range		
	0.9-3 kb. Library is normalized and enriched for		
	full-length clones and was constructed by C. Gruber		
	(Invitrogen). Research Genetics tracking code 013. Note:		
	this is a NIH MGC Library."		
BASE COUNT	256 a	154 c	227 g
ORIGIN		160 t	
Query Match	80.9%	Score 723;	DB 10; Length 817;
Best Local Similarity	97.7%	Pred. No. 1.6e-130;	
Matches 765;	Conservative 0;	Mismatches 15;	Indels 3; Gaps 3;
0Y	4	agtggagtgagggtgaacaagaatggcgagccgagaggtggagctccataagctaaagctt	63
DB	35	AGTGAGTGTGAGGGGTGAACAAGATGGCGAGCCGAGAGCGTGGAGCTCATTAAGCTT	94
0Y	64	gccgaactaagaagaatgtctgtcgtcgtgtgttggagaccaaggaataagaagcaat	123
DB	95	GCCGAACATAAACAAGATGCTTCTGCTGCTGCTTGGAGACCACAAGGAATAAACAAGAT	154
0Y	124	ctttacccaagctccagagcatatttgaagaacatgtctgaagagaggaagcaatgaaga	183
DB	155	CTTTATCCACAGACTCCAGCATATCTTGAAGAACATGCTGAAGAGAGCAATGAAGAA	214
0Y	184	gattgactgtggagatgaacaacagaggaagaagaacaacaaagccatgtagctcccttcaa	243
DB	215	GATTGTACTGTGGGAGATGAACAACAGAGGAAGAAGAAACAAGCCATTGAGCTCCCTTCAA	274
0Y	244	gaggaagaaccccttgaaaaacgttgcattgtgtgcagcagagagaagaagtgtgtaaat	303

Db	275	GAGGAAGAACCCCTGAAAAAACTGTTCATGTGGCAGCAGAGAGAAAGTGCTGAAATTT	334
QY	304	acatctgcaataaccacagactgagaaatgcaagaagagcgctgaaacgaattcaatgaact	363
Db	335	ACATCTGAAATACACAGACTGAGGATTCGAGAGAGCGCTGAACGATTCGAATGTA	394
QY	364	gtgaactgtagaagtaagaagaagctgctcgggcagctaggtttggaattcttcagttcca	423
Db	395	GTGACTCTGGAGAGTAAGAAAGCTGCTCGGCGACCTAGGATTTGGGATTTCTTCAGTTCCA	454
QY	424	acaaaagttctgtctctgatacaaaacctatggttaacttggataagctgaaggaaga	483
Db	455	ACAAAAAGCTGTGCTGATATACAAACCTATGTTAACTTGATTAACCTGAAGGAAAGA	514
QY	484	gtccaagaattggtttgaatgctcttcctccaatctccagaagctgtaagatgataga	543
Db	515	GCTCAAGATTTGGCTTTGATATCTCTTCAATCTCCAAAGTCTGAAAGATGATGAGAA	574
QY	544	ctgaagaagaaggaagagcgatcttggatctgcaagaattcaagctggaactggaacca	603
Db	575	CTGAAAAAGAGAGAGAGCGATTTGGATTTGTCACAAATTCAGCTGGAATCGAAACCA	634
QY	604	gaggaatacagagcagaagaagaagaagcagaagcgctttggatt-gcctatgaaa	662
Db	635	GAGGATACAGAGCAAGAAAGAGAAAGAGAGAGCGCTTTGGATTTGGCTCATGTAAG	694
QY	663	agttccatgaattctgtctctccagtgctttccaattctctcctctctcttctgtgaca	721
Db	695	AGTTCTGTAATCTTCTGTCTCTCCAGTGTTCATTTCTCTCTCTCTCTCTCTCTCTCT	754
QY	722	tatatgacctaaatgcacagtcatactgccc-tacgtctcgtccatgaaatgaaagcatgta	780
Db	755	TATATGCGCTAAATWTCAGTCATGTGCTTACGTCTCTCGTCGATGAGGAGCATTTGTA	814
QY	781	ccc 783	
Db	815	ccc 817	
RESULT	12		
LOCUS	BG493148		
DEFINITION	6025541829p1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4672974 5',	744 bp	linear EST 27-MAR-2001
ACCESSION	BG493148		
VERSION	BG493148.1	GI:13454660	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 744)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: rsraus@biml.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: CLONTECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.llnl.gov		
	Plate: L10C1490 row: 1 column: 07		
	High quality sequence stop: 731.		
FEATURES	Location/Qualifiers		
source	1..744		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4672974"		
	/clone_id="NIH_MGC_59"		
	/tissue_type="mucoepidermoid carcinoma"		

/lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
 SfiI (ggccgcccggcc); Site_2: SfiI (ggccattggcc):
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGGCATATGACC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCCGCGAGCGGCGCATAGCTAAAGCT-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA). Note: this is a NIH-MGC
 Library."

BASE COUNT 248 a 131 c 200 g 164 t 1 others
 ORIGIN

Query Match 80.1%; Score 716; DB 10; Length 744;
 Best Local Similarity 98.4%; Pred. No. 3.8e-129;
 Matches 733; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

OY 3 gagtggagtgagggttaacagaatgagccgagagcgggtgagctccataaactaaagct 62
 |||||||
 Db 1 GAGTGGAGTGAAGGGTAAACAGATGGCGACGAGGGGAGCTCCATAGCTAAAGCT 60
 OY 63 tgcgaactaaagcaagaatgtctgtcgtgtgttggagacaaagggaataaacaaga 122
 |||||||
 Db 61 TGCCGAACTAAAGCAAGATGTCTTGCTGCTGTTGGAGACCAAGGGAAATTAACCAACA 120
 OY 123 tctatccacagactccagagcatcttgaagaacatgctgaaggaggaacaaatgaaga 182
 |||||||
 Db 121 TCTTATCCACAGACTCCAGGCAATCTTGAAACATGCTGAAGAGGCAAAATGAACA 180
 OY 183 agatgactggagagtgaacagaagaagaacaaagccattgaactccctctgca 242
 |||||||
 Db 181 AGATTTACTGGAGATGAACAGAGAGAGAGAAACAAAGCCATTGAGCTCCTGTCA 240
 OY 243 agaggaagaacccccctgaaaaaacctgtgaltggtgcagcagaagaagaagtgtaaaat 302
 |||||||
 Db 241 AGAGGAGAGACCCCTGAAAAAACTGTGATGTGCGACAGAGAGAAAGTGTGTAAT 300
 OY 303 tacatctgaataacacacagactgtagaagaatgagaagggtctgaacgttaaatgtacc 362
 |||||||
 Db 301 TACATCTGAAATATCCACAGACTGAGAGAAATGCAAGAGGCTGAACGATTAATGTACC 360
 OY 363 tctgagcttgagagtaagaagaagctgcgcgagcaagctgaagtttggaattcttcagttcc 422
 |||||||
 Db 361 TGTGAGCTTGGAGATTAAGAAAGCTGCTCGGACAGTAAAGTTTCTTCAGTTCC 420
 OY 423 aacaaaagctctgcatctgatacaaacctaatgtaacttgatgaatgaagaaag 482
 |||||||
 Db 421 AACAAAGGTCGTGATCTGATTAACAAACCTATGTTAATCTTGATTAACCTGAAGAAAG 480
 OY 483 agctcaagaattgttgaagtcttccaatctccagaagaagctcgaagaatgagagaa 542
 |||||||
 Db 481 AGCTTAAAGATTGTGTGAATGTCTTCAATCTCCAAAGAGTCTGAAGATGATGAGAA 540
 OY 543 actgaaaaaagaagaagagcgaatttggaatgtaacaagtctcagctggaactggaaccac 602
 |||||||
 Db 541 ACTGAAAAAGAGAGAGACCGA-TTGGGATTTGTCAAGTTAGCTGGAACCTGGAACCAAC 599
 OY 603 agagatacagaagcaagaagaagaagaagcagagcagcttggagattgacctgtgaagaa 662
 |||||||
 Db 600 AGAGATACAGAGCAAG 659
 OY 663 agttcctgaactctctgtctccagtggttccatctctcctctctctgtgtcaat 722
 |||||||
 Db 660 AGTTCTCTGACTTCTGTCTCCAGTGTTCATTTCTCTCTCTCTCTGCTGACAT 719
 OY 723 atatgctcaaatgacagtcagtgtg 747
 |||||||
 Db 720 ATATGCTCAATGACACAGTCCATGTG 744

RESULT 13
 BG720065 781 bp mRNA linear EST 08-MAY-2001
 LOCUS 602691562f1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4823943 5',
 DEFINITION mRNA sequence.

ACCESSION BG720065
 VERSION BG720065.1 GI:13999252
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 781)
 NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki
 Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov

Plate: LHAM10734 row: k column: 16
 High quality sequence stop: 765.

FEATURES
 source
 Location/Qualifiers
 1..781

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4823943"
 /clone_1ib="NIH_MGC_97"
 /lab_host="DH10B"

/note="Organ: testis; Vector: pBluescriptR (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
 size-selected for average insert size 2.2 kb and
 normalized to R0F 5. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH-MGC Library."

BASE COUNT 256 a 142 c 210 g 173 t
 ORIGIN

Query Match 79.9%; Score 714.6; DB 10; Length 781;
 Best Local Similarity 98.7%; Pred. No. 6.9e-129;
 Matches 773; Conservative 0; Mismatches 4; Indels 6; Gaps 5;

OY 2 ggaatggagatggaggttaacagaatgagccgagagcgggtgagctccataaagctaaagc 61
 |||||||
 Db 4 GGAGTGGAGATGAGGGGTAAACAAGATGGCGACCGAAGCGGTGAAGCTCCATTAAGCTTAAAGC 63
 OY 62 ttgccgaactaaagcaagaatgtctgtcgtgtgttggagacaaagggaatcaagcaag 121
 |||||||
 Db 64 TTGCCGAACATAAGCAAGATGTCTTGCTGCTG- TTGGAGGCCAAAGGAGATTAAGCAAG 122
 OY 122 atctatccacagactccagacatcttgaagaacatgtctgaagaaggaggaatgaagaa 181
 |||||||
 Db 123 ATCTTATCCACAGACTCCAGGCAATATCTTGAAGAACATGCTGAAGAGAGCAAAATGAAG 182
 OY 182 aagatgtactggagatgaacagaggaagaagaacaaagccatgtgagctccctgtca 241
 |||||||
 Db 183 AAGATGTACTGGAGATGAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
 OY 242 aagaggaagaacccccctgaaaaaactgttgatgtgtgagcagaagaagaagtggtgaaga 301
 |||||||
 Db 243 AAGAGAGAGAACCCCTGAAAAAACTGTTGATGTGGCAGACAGAGAGAGAGAGAGAGAGAG 302

QY 302 ttacatctgaataccacagactgagaaatgcagaaagggctgaacgattcaatgtac 361
|||||
Db 303 TTACATCTGAATAATACACAGACTGAGAAATGACAGAGAGGCTGAAAGATTCAATGTAC 362
QY 362 ctggaacttgaggagtaagaagaactcgtcgggagcgtgaagtttgggaattcttcagtc 421
|||||
Db 363 CTGTGACCTTGGAGAGTAAGAAAGCTCTCGGACGCTAGGTTTGGGATTTCTTCAGATTC 422
QY 422 caaaaaaagctctgcatctatatacaaacctatggttaacttgatagaagtgaa 481
|||||
Db 423 CAATAAAGGCTCTCATCTCATATACAAACCTATGGTTAACTTGATTAAGCTGAAGGAAA 482
QY 482 gagctcaagaattggttgtaagtctcttcaatctccagaagaatctgaagatgataga 541
|||||
Db 483 GAGCTCAAGAATTGTTGTAATGTCTCTTCAATCTCCAGAAAGTCTGAAGATGATGAGA 542
QY 542 aactgaaaaagagagagagagatttggattgtcacaagttcagctggaactggaacca 601
|||||
Db 543 AACTGAAAAAGAGAGAGAGAGGATTTGGATTGTCACAAAGTTCAGCTGGAACCTGGAACCA 602
QY 602 cagaagatacagagagcaaaagagagagagagagagagagagagagagagagagagag 661
|||||
Db 603 CAGAGGATACAG -GGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659
QY 662 aagttcctgactcttctgtctcagtggtttcca-tttctcctctcttcttctgtac 720
|||||
Db 660 AAGTTCCTGATCTTCTGTTCTCCAGTGTTCCTCAATTTCTCTCTCTCTCTCTGTTAC 719
QY 721 atatactcctaataatgcacagtcagtctcagtcagtcagtcagtcagtcagtcagtcag 780
|||||
Db 720 ATATATCTCAATAATGACACAGTCATGCTCCTACGTCG-CTCGCAATGAGGAGCATGTA 778
QY 781 CCC 783
N1
Db 779 CCC 781

RESULT 14
LOCUS BG387900 974 bp mRNA linear EST 12-MAR-2001
DEFINITION 602412920F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4521268 5',
mRNA sequence.
ACCESSION BG387900
VERSION BG387900.1 GI:13281346
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 974)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
[found through the I.M.A.G.E. Consortium/LLNL at:](http://image.llnl.gov)
<http://image.llnl.gov>
Plate: LLAM10420 row: h column: 05
High quality sequence stop: 759.
Location/Qualifiers
1..974
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4521268"
/clone_1ib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-Sport6; Site_1: NCI;

Site_2: SalI; Cloned unidirectionally; Oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library.

BASE COUNT 326 a 218 c 245 g 185 t
ORIGIN

Query Match 79.6% Score 711.8 DB 10; Length 974;
Best Local Similarity 98.1% Pred. No. 2.3e-128;
Matches 741; Conservative 0; Mismatches 12; Indels 2; Gaps 2;

QY 19 aacaagatgagcagagagagcgttgagctccataagctaaagcttgcgcgaactaaagaa 78
Db 1 AACCAAGATGGGACCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 79 gaatgtctgtcgttggttgagagagagagagagagagagagagagagagagagagag 138
Db 61 GAATGTCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 139 cagcagatctcttgaagaaacatgctgaaagagagagagagagagagagagagagagag 198
Db 121 CAGGATATCTTGAAGAGACATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 180
QY 199 gaacag 258
Db 181 GAACAG 240
QY 259 gaaaaaactgttgatgt 318
Db 241 GAAAAAACTGT 300
QY 319 cagactgag 378
Db 301 CAGACTGAG 360
QY 379 aagaagctgtcgt 438
Db 361 AAGAAGCTGCTCGGAG 420
QY 439 tctgatacaaacctatgtgttaacttgaataagctgaaagagagagagagagagagag 498
Db 421 TCTGATTAACAACCTATGCTTAACTTGATTAAGCTGAAGAGAGAGAGAGAGAGAGAG 480
QY 499 ttgaatgtcttcaatctcaaaagagagagagagagagagagagagagagagagagag 558
Db 481 TTGAATGTCTTCAATCTCCAGAAAGTCTGAAGATGATGAAGAACTGAAAAAGAGAGAG 540
QY 559 gagcagatttgagatgtgtacaaagttcagctgaaactgaaacacagagagatcacagagca 618
Db 541 GAGGATTTGGGATTTGTCAACAGTTCAAGCTGGAAGCTGGAACACAGAGAGATACAGAGCA 600
QY 619 aagaagagga-aaagagcagagcgtttggagattgctgtatgaaagttcctgtacttt 677
Db 601 AAGAAGAGAGCAACAG 660
QY 678 ctgttctcagatgttttccattcttctcctctcttcttctgtgtacatatatgcttaatgca 737
Db 661 CTGTTCTCCAGATGTTTTCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 719
QY 738 cagtcagtgtgctcagtcctgtcctcctcagatgagg 772
Db 720 CAGTCATGTGCCCTACAGTCGTGCTCTGCAATGGAGG 754

FEATURES
source
1..974
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4521268"
/clone_1ib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-Sport6; Site_1: NCI;

RESULT 15
LOCUS BG289677 1139 bp mRNA linear EST 21-FEB-2001
DEFINITION 60238464F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4513499 5',
mRNA sequence.
ACCESSION BG289677
VERSION BG289677.1 GI:13045679
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LLM10400 row: d column: 12
High quality sequence stop: 697.
Location/Qualifiers
1. 1139
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4513498"
/clone_lib="NIH_MGC_93"
/tissue_type="transitional cell papilloma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: bladder; Vector: PCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: This is a NIH_MGC Library."
BASE COUNT 380 a 224 c 315 g 219 t 1 others
ORIGIN

Query Match 78.9%; Score 705.8; DB 10; Length 1139;
Best Local Similarity 94.7%; Pred. No. 3.3e-127;
Matches 785; Conservative 0; Mismatches 37; Indels 7; Gaps 5;

QY 14 ggggttaacaagaatgagcgaccgagcgtgtgagctccataagcttaagcttcggaactaa 73
DB 1 GGGGTAAACAAGATGGCGACCGAGCGGTGAGCTCCATTAAGCTTAAGCTTGGCACTTA 60
QY 74 agcaagaatgtctgtcgtgtgtgttgagacccaaggaaataagcaagaattcattccaca 133
DB 61 AGCAAGAATGTCTGCTCGTGTGTTGAGACCAAGGAAATAAGCAAGATCTTATCCACA 120
QY 134 gactccaggcatatttgaagaacatgtctgaagaaggagcaaatgaagaagatgtactgg 193
DB 121 GACTCCAGGCATATCTTGAAGAACATGCTGAAGAGAGGCAATGAAGAGATGTACTGG 180
QY 194 gagatgaacaagaagaagaagaagaagaagccattgagctccctgtcaagaaggaaagac 253
DB 181 GAGATGAACAAG 240
QY 254 cccctgaaaaaaactgtgatgt 313
DB 241 CCCCTGAAAAAACTGT 300
QY 314 taacacagaactgtgaagaatgtgaagaagggtgtgaagatcaatgtactgtgagcttgg 373
DB 301 TACCACAGACTGAGAGAAAGCAGAAAGAGGCTGAAGCATTAATGATGATGAGCTTGG 360
QY 374 agagtaagaagaactgtcgtgagcagctaggtttgagattctcagttccacaagaagctc 433
DB 361 AAGATGAAGAAAGCTGCTGGGCAAGCTAGGTTGGGATTTCTTCAGTTCCAAACAAAGGTC 420
QY 434 tglcatctgatatacaaaactatgttgaacttgataagctgaaggaaagaagctcaaat 493
DB 421 TGTCAATGTGATTAACAACCTATGTTAATTGATTAAGCTGAAGGAAAGAGCTCAAGAT 480
QY 494 ttggtttgaatgtctcttcaatctccagaagcttgaagaatgaagaactgaagaaga 553

DB 481 TTGGTTGAATGTCCTTCATCTCCAGAAAGCTGAAGATGATGAGAAACTGAAAAAGA 540
QY 554 ggaagagcgaatttgatgttcacaaagtccagctggaactggaaccagagagatag 613
DB 541 GGAAGAGCGATTTGGGATTTGCACAAAGTTCAAGCTGGAACCTGAAACACAGAGATACG 600
QY 614 aggcaagaagaagaagaagaagaagagcgccttggagattgcctgatgaaaagttcctgata 673
DB 601 AGCAAGAAG 657
QY 674 cttctgtctccagtggttttccattctctcc-ttctctgtgtcacaataatgcttaa 732
DB 658 CTTTCTGTCTCCAGTGTGTTTCCATTTCTCTTCTTCTTGTGTCACATATATTCCTTA 717
QY 733 atgcacagtcagtgcctcactgcctgcctgcctgcctgcctgcctgcctgcctgcctgcct 789
DB 718 AATGCCAGCATGTGCTACGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
QY 790 catcatgaactgcggcagcagcttgaactaattgtatgtgttgcagctttaa 838
DB 778 ACTCCCTTACTGCGGCGCGCTTGGCTTAATTTGCGGGTTCAACGCTTA 826

Search completed: June 19, 2002, 13:36:37
Job time: 6345 sec